Wax

ARPA Network Information Center
Stanford Research Institute
Menlo Park, California 94025
NIC #30798
Packet Radio Remporary Note #110

Network Analysis Corp.
May 1974

ON THE CAPACITY OF SLOTTED ALOHA NETWORKS AND SOME DESIGN PROBLEMS*

by

Israel Gitman

Network Analysis Corporation
Beechwood, Old Tappan Road
Glen Cove, New York 11542

ABSTRACT

Packet Switching over radio channels with random access schemes is of current interest for local distribution systems and for satellite channels. This mode of operation is useful when the communicating devices are mobile and when the ratio of the peak to average data rate requirement of each device is high. Such systems have been analyzed for the case in which all communicating devices are within an effective transmission range of each other; either directly or through the satellite.

In this paper, we address broadcast networks in which originating devices cannot directly reach the destination receiver. Thus, devices are introduced which receive these packets and repeat them to the destination. The capacity (maximum throughput) of such systems is determined, and design problems related to the number of repeating devices and the usefulness of directional antennas are resolved.

---

* This work was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC 15-73-C0135.

1. INTRODUCTION

One way to categorize channel allocation schemes for data transmission is the following:

(i) Fixed assignment (FDM, TDMA)

(ii) Dynamic assignment with centralized control (polling schemes, reservation upon request)

(iii) Dynamic assignment with distributed control (loop networks, random reservation schemes)

(iv) No assignment (random access schemes)

It has been recognized that a fixed allocation of channel capacity is extremely wasteful when the traffic of users is of a bursty nature; that is, when the traffic requirements of the users can be characterized as having a high peak to average data rate. Users can be so characterized in an inquiry response application. In fact, if one characterizes a set of users by the number and by the ratio of the peak to average data rate requirements of each user, one can conjecture that when the above number and ratio are increasing, one obtains a higher channel utilization when proceeding (along the categorization) from the fixed assignment to the no assignment allocation schemes. For example, when the time delay to make a reservation or the average time between two consequtive pollings of a user is large compared to the fraction of time that the user wishes to use the channel, then the dynamic assignment with centralized control, becomes inefficient (apart from the need for a system for polling or making reservations).

Roberts [7] has demonstrated the cost advantages of a random reservation scheme and a random access scheme over fixed assignment, when the number of users increases and the average traffic requirement

per user is kept constant. A somewhat more formal justification for sharing a channel was given by Kleinrock and Lam [4], we quote: "Rather than provide channels on a user-pair basis, we much prefer to provide a single high-speed channel to a large number of users which can be shared in some fashion; this when allows us to take advantage of the powerful 'large number laws' which state that with very high probability, the demand at any instant will be approximately equal to the sum of the average demands of that population." Gitman, Van Slyke, and Frank [3], have addressed the problem of splitting a channel between two classes of users. It was shown that in almost all cases, sharing the channel results in a higher utilization of the total channel capacity.

In this paper, we consider a packet switching network in which a single radio channel is shared by all communicating devices. Devices access the channel using the so called "slotted ALOHA" random access scheme. When a random access scheme is used, there is a possibility that more than one packet are simultaneously received by a receiver due to independent transmissions of several devices. In that event, it is assumed that none of the packets is correctly received, and the corresponding devices have to retransmit their packets. One can see that the number of packets transmitted is larger than the number of originating packets. That is, part of the channel capacity is used up by the wasteful collisions and are not considered as effective channel utilization since it does not contribute to the throughput. For example, if each packet is transmitted two times, on the average, before it is successfully received, then the maximum effective utilization of the channel (or the effective channel capacity) is one-half of the given capacity, the other one-half is used up by the non-successful transmissions. The first problem that one faces is to determine the maximum effective utilization (or system capacity) that can be obtained. This is one of the problems addressed in the paper.

If the channel is offered a higher rate of traffic than its effective capacity, the system becomes unstable in the sense that the number of transmissions increases with time and the throughput decreases with time until zero throughput is obtained. This implies another problem in random access schemes and that is, the control of the offered rate and retransmission strategies so as to obtain the highest possible channel utilization [5], [8]. It is clear however, that retransmissions have to be randomized in time since otherwise once a collision between two devices occurred, it will persist.

The first packet radio channel system where devices use a random access scheme, known to us, was analyzed and implemented at the University of Hawaii [1]. The random access scheme used is the so called "pure" or "unslotted" ALOHA. In this scheme, every terminal transmits its packets independent of any other terminal or any specific time. That is, the terminal transmits the whole packet at a random point in time; the terminal then times out for receiving an acknowledgement. If an acknowledgement is not received, it is assumed that a collision occurred and the packet is retransmitted after an additional random waiting time. Abramson [1] obtained that the capacity (effective) of the channel is 1/2e of the given capacity when the number of terminals is very large and when the point process of the beginning of packet transmission onto the channel is Poisson.

It was realized that a gain in capacity can be obtained if the channel was slotted into segments of time whose duration is equal to the packet transmission time, and when terminals are required to begin the transmission at the beginning of a time slot. The access scheme is random in the sense that terminals transmit into a random slot in time and retransmit after waiting a random number of slots. This scheme is called "slotted ALOHA." Roberts [6] has shown that the capacity of this scheme is 1/e of the given capacity, using the same

assumptions as Abramson. The slotted ALOHA random access scheme was further analyzed in [2], [3], [4], and [5] for the case of a small number of terminals and when there is a mixture of traffic from a small number of "big" users and a large number of "small" users.

Digitized by the Internet Archive
in 2025 with funding from
Amateur Radio Digital Communications, Grant 151

https://archive.org/details/temporary-note-110

2. PROBLEM DESCRIPTION

In a network context, the analyses of all the references address a "single hop network." That is, when considering a terrestrial system, it is implicitly assumed that all devices are within an effective distance of each other; the same is true when considering a satellite channel, since the satellite echoes the packets to the destination station. Such a network can be described as one in which there is a single receiver and many transmitters, all of which are within an effective transmission range to the receiver, and to each other.

In cases in which the transmission range of terminals is not sufficient to reach the destination receiver, it is necessary to introduce another device which will receive the packets from the terminals and repeat them to the final destination. We shall call this device a "Repeater." Thus, one obtains a 2-Hop network, one hop from terminals to repeaters, and second hop to the final receiver, which we shall call a "Station."* Such a network can be used for local distribution and collection of traffic, in which case the station is a gateway to a point-to-point network. We particularly consider a model in which there is a large number of terminals in the neighborhood of each repeater and that the transmission range of terminals is short so that a terminal can reach only one repeater, as shown in Figure 1. Our model can be useful as a distribution model for a suburban area, where instead of supplying each terminal with a powerful transmitter, we allocate repeaters which collect the traffic from terminals.

In a military application, one can think of a large unit which uses such a 2-Hop network, where each subunit(the terminals of which are relatively close to each other and may be mobile) has its own repeater, and the station is at the Headquarters. In fact, the network can be operative when the whole unit is moving, providing each submit carries its repeater or station along with it, e.g., a fleet of ships.

---

* At the University of Hawaii, one of the campuses on an island was out of range and a simple repeater was recently introduced.

This network can also model a satellite case as shown in Figure 2. In the analyses of a satellite channel reported, it was assumed that each terminal is a ground station which originates its traffic or which is connected to a point-to-point network. However, suppose there are clusters of (possibly mobile) terminals, where in each cluster, they are relatively close to each other, and wish to communicate with a remote central computer installation. Then, one can devise a Ground Station Repeater to communicate with the terminals and the ground station repeater may use a satellite channel to transmit (second hop) to the computer installation.

Before we specify the problems to be addressed, we comment on the operation of the repeater. We have indicated that a terminal retransmits its packet after a time out if it does not receive an acknowledgement. The repeater can operate in a very simple manner in which it repeats only once a packet that it correctly receives. In this case, the terminal has to time out for a longer period of time to wait for an acknowledtement from the station (end-to-end ack). Alternatively, the repeater can be made responsible for the successful transmission on the second hop, by acknowledging the terminal, storing the packet, and transmitting it until it receives an acknowledgement from the station (hop-by-hop acknowledgement). This problem has been considered in [9] where it was shown that a hop-by-hop acknowledgement operation is more efficient. Thus, we will assume a repeater of that type.

The first problem that we address is to determine the network capacity as a function of the number of repeaters and the interference between repeaters. Another problem of interest, is to determine the capacity bottleneck (critical hop); i.e., whether the capacity bottleneck is on the hop from terminals to repeaters or on the hop from repeaters to station. The system is assumed to have two channels; one channel is used for transmission from terminals via repeaters to

the station and the second channel for transmission from the station via repeaters to the terminals. The traffic of acknowledgement packets is not considered; which implicitly assumes that there is a separate channel for acknowledgements. The two channels are analyzed separately and design questions such as the following are considered: Is it useful to have directional antennas at repeaters when transmitting to the station ? This question is relevant in the terrestrial system, since in the satellite system, the ground station repeaters will presumably be out of range from each other, and use directional antennas. Similarly, in a terrestrial system, one may ask about the usefulness of using a directional antenna at the station when transmitting to repeaters. Other questions relate to the possibility of using several transmitters and antennas at the station when transmitting to repeaters.

3. TRANSMISSION FROM TERMINALS TO STATION

Consider a system where m repeaters receive packets from terminals and repeat the packets to a single station, as shown in Figure 3. We donote by G and S the rate of packet transmission per slot and the rate of successful packet transmission per slot, respectively. Specifically, let G_{1i} and S_{1i} be the rates of transmission from terminals to repeater i, and G_{2i} and S_{2i} the rates from repeater i tc the station. We wish to obtain the probability that a repeater is idle.

A single hop network is the case in which a set of terminals transmit to a repeater and the repeater is the final destination and does not repeat the packets. Thus, the probability that a transmission from a terminal to the repeater is successful is the probability that no other terminal transmits into that same slot. That is, if we assume that a packet is transmitted into the first slot after it becomes ready for transmission then the probability of success is the probability that no new packet has arrived and that no other packet has been scheduled for retransmission in the interval of time of the preceeding slot. In the network case however, a transmission from a terminal to a repeater will not be successful also in the case when the repeater uses the same slot for transmitting to the station, or if another repeater, within an effective transmission range of the first and which uses an omnidirectional antenna, uses the same slot for transmission to the station.

Throughout the paper, we use the following assumptions. The combined point process of packet origination and packets schedulled for retransmission, from each set of terminals to a repeater, is Poisson. Thus, the probability of no arrival during a slot time τ is $e^{-G_{1i}\tau}$; we use $\tau \equiv 1$. The probabilities of transmission by a repeater into different slots are independent. The probability of

transmission by two or more repeaters into a randomly chosen slot are mutually independent; and the probability of transmission into a random slot by a terminal and by a repeater are independent. The assumption of a Poisson distribution for the process of packet origination plus retransmission has been questioned in previous publications. The validity of this assumption is important when one considers packet delays, since if one assumes a Poisson point process for packet originations, then the assumption that the combined process, of origination plus retransmissions, is Poisson is valid only when one allows very large packet delays. In this paper, we do not require finite delays and our interest is in the ultimate system capacity.

Finally, let Q_i denote the set of repeaters which have an effective transmission range to repeater i; Q_i includes repeater i. Under these assumptions we can write:

$$P_r[\text{repeater i is idle}] = e^{-G_{1i}} \prod_{j \in Q_i} (1-G_{2j}) \qquad (1)$$

Similarly, the probability that the station is idle can be written as:

$$P_r[\text{station is idle}] = \prod_{j=1}^{m} (1-G_{2j}) \qquad (2)$$

We now make a few assumptions which simplity the computation but enable us to answer the questions of interest. Specifically, we assume that $G_{1i} = G_1$, $S_{1i} = S_1/m$ for all i, that repeaters share equally the load so that $G_{2i} = G_2/m$ and $S_{2i} = S_2/m$ for all i, and that all repeaters have the same number of interfering repeaters, $n(Q_i) = I$ for all i. We refer to I as the interference level.

We can now write the throughput equations for hop 1 and hop 2:

$$S_1 = \sum_{1}^{m} G_1 \left(1-\frac{G_2}{m}\right)^{I} e^{-G_1} = mG_1 \left(1-\frac{G_2}{m}\right)^{I} e^{-G_1} \qquad (3)$$

$$S_2 = \sum_{1}^{m} \frac{G_2}{m} \left(1-\frac{G_2}{m}\right)^{m-1} = G_2 \left(1-\frac{G_2}{m}\right)^{m-1} \qquad (4)$$

Note that when m=1 all transmissions from the repeater to the station are successful since it does not interfere with its own transmissions.

If a repeater were a traffic source and a traffic sink, then G_1 and G_2 could have been considered as independent variables. In our case, however, the intensity of the processes on the two hops are related. In particular, we consider traffic rates in which the system can operate at steady state. That is, if one observes the system for a long period of time, then all the packets which successfully arrive to repeaters also successfully arrive to the station. Thus we can use the conservation law at repeaters, namely $S_1 = S_2$. This results in the relation:

$$G_2 = mG_1 \left(1-\frac{G_2}{m}\right)^{I+1-m} e^{-G_1} \qquad (5)$$

One can now study the system performance as a function of parameters m and I, with one independent variable.

A. Complete Interference System, I = m

This case is applicable to terrestrial networks in which repeaters are either placed relatively close to each other or use powerful transmitters; either of which results in the interference among all repeaters. For this case we obtain.

$$G_2 = \frac{m\, G_1\, e^{-G_1}}{1+G_1 e^{-G_1}} \qquad (6)$$

and

$$S_1 = \frac{m\, G_1\, e^{-G_1}}{(1+G_1\, e^{-G_1})^m} \qquad (7)$$

The capacity of this system is given by the maximum of S_1.

$$\frac{d\, S_1}{d\, G_1} = \frac{m\, e^{-G_1}}{(1+G_1\, e^{-G_1})^{m+1}} (1-G_1)[1-(m-1)\, G_1\, e^{-G_1}] = 0 \qquad (8)$$

By examining (8) one finds that there is one stationary point, a maximum at $G_1 = 1$, when $m<4$. When $m\geq 4$ there are three stationary points; a minimum at $G_1 = 1$, and two maximum points of the same value at G which satisfies $1-(m-1)\, G\, e^{-G} = 0$. Substituting these values into (7), one obtains the capacity of this (complete interference) network as a function of m:

$$\text{Network Capacity} = S_1^* = \begin{cases} \dfrac{m}{e\,(1+\frac{1}{e})^m} & m<4 \\ \\ (1-\dfrac{1}{m})^{m-1} & m\geq 4 \end{cases} \qquad (9)$$

Notice that the network capacity is lower than the capacity of a single hop network , 1/e, when m=1; it is higher than 1/e for m>1, and tends to 1/e in the limit when $m\to\infty$.

Figure 4 shows the network throughput S_1 as a function of the rate of transmission from terminals to a single repeater. It is interesting to observe the rate of change of S with respect to G. For example, when m is large (e.g. m=8) then the system becomes more sensitive to variations in the value of G. On the other hand, there are values of m, for example, m = 2 or 3 in the complete interference case, for which the maximum throughput region is quite flat.

B. The Critical Hop

To determine the critical hop, one has to obtain the capacity of the two hops of the network. The capacity of the hop from repeaters to the station is independent of G_1 and I (see Eg. (4)), and is given by $(1-1/m)^{m-1}$. The capacity of the hop from terminals to repeaters, on the other hand, depends on G_2 and I. Note that $0 \leq G_2 \leq m$; $G_2 > m$ is not realizable since we assume that a repeater has one transmitter and cannot transmit more than one packet per slot. For any G_2 in the above range, the capacity of hop 1 (see Eg. (3)) increases with m.

Thus, there exists an m_o for which the capacity is higher than that on the hop from repeaters to station, and the latter becomes the critical hop. Furthermore, m_o depends on I, and for $I_2 \geq I_1$, $m_o(I_2) \geq m_o(I_1)$. Thus for $m > m_o$ the critical hop is that from repeaters to station, and for $m \leq m_o$ the critical hop is from terminals to repeaters. For example, for I=m (see Eg (9)) $m_o = 4$, and for I = m-1, $m_o = 3$.

Perhaps a more direct way to answer the question of the critical hop is to obtain the system capacity as a function of m and I. Then, whenever the capacity is smaller then $(1-1/m)^{m-1}$ the critical hop is that from terminals to repeaters and when the capacity equals this expression the critical hop is from repeaters to station. However, it is difficult to obtain a closed form solution in the general case.

C. Number of Repeaters and Directional Antennas at Repeaters

The considerations of directional antennas at repeaters apply only to the terrestrial system. In the satellite system, ground station repeaters will use directional antennas when transmitting to the satellite and omnidirectional antennas when communicating with terminals.

The effect of directional antennas at repeaters in the terrestrial system is that the transmission from repeaters to the station is directed towards the station and does not interfere with the transmission of terminals to other repeaters. Thus, it is the special case with I=1. We notice, however, that directional antennas do not increase the capacity of the hop from repeaters to station because all antennas are directed towards the same physical location where the station is placed and where the conflicts may occur.

Figure 5 shows the capacity of the system as a function of m, for I=m and I=1, which is equivalent to omnidirectional and directional antennas (or satellite system) respectively. One can see that there is a gain in capacity when using directional antennas only when m=2, and a small gain for m=3. For $m \geq 4$ the capacity does not increase because the critical hop is between the repeaters and the station, so that it does not matter how much one can get through from terminals to repeaters.

As far as the number of repeaters is concerned, one can see from Figure 5 that the maximum system capacity is obtained when m=2 in the non-interference case and when m=3 in the complete interference case. Thus 2 or 3 repeaters would be a good design; any additional repeaters that may be added because of other considerations (such as area coverage) will result in a reduction in the system capacity.

4. TRANSMISSION FROM STATION TO TERMINALS

In this section, we consider the second channel which is used for transmission from the station to terminals via repeaters. In the terrestrial system, it is assumed that the effective transmission range of the station is such that it interferes with the transmission from repeaters to terminals, as shown in Figure 6. However, we assume that terminals cannot directly receive from the station or from the satellite (otherwise, it becomes a single hop network and the capacity is 1). We use the notation shown in Figure 6, where the first hop is that from the station to repeaters and the second hop is from repeaters to terminals.

We use similar assumptions to the ones made in the previous section. Specifically, we assume that the probabilities of transmission by a repeater into different time slots are independent, that the probability of transmission of two or more repeaters into a randomly chosen slot are mutually independent, and that the probability of transmission by the station and by repeaters into a random slot are mutually independent. Finally, we simplify by assuming that repeaters share equally the load; by which we mean that $S_{2i} = S_2/n$ and $G_{2i} = G_2/m$, for all i. The equations which relate the rate of transmission to the rate of successful transmission on the two hops can now be written:

$$S_2 = \sum_{1}^{m} \frac{G_2}{m} (1 - G_1) (1 - \frac{G_2}{m})^{I-1} = G_2 (1 - G_1) (1 - \frac{G_2}{m})^{I-1} \quad (10)$$

$$S_1 = G_1 (1 - \frac{G_2}{m})^{I} \quad (11)$$

I is the interference level as in the previous section.

We assume again steady state system operation and use the conservation law at repeaters, i.e., $S_1 = S_2$. This results in

$$G_2 = \frac{G_1}{1 - G_1 \frac{m-1}{m}} \tag{12}$$

One can now substitute (12) into (10) and (11) to obtain the throughput as a function of one independent variable. For S_1, one obtains:

$$S_1 = G_1 \left(\frac{1 - G_1}{1 - G_1 \frac{m-1}{m}} \right)^I \tag{13}$$

To obtain the capacity one can maximize either S_1 or S_2. Equating dS_1/dG_1 to zero, we obtain that there are three stationary points. A minimum at $G_1^1 = 1$ and two maximum values at

$$G_1^{2,3} = \frac{(2m + I - 1) \pm \sqrt{4mI + (I-1)^2}}{2(m-1)} \tag{14}$$

We now examine the constraints. For the system to be realizable,

$$0 \leq G_1 \leq 1 \; ; \; 0 \leq G_2 \leq m \; ; \; 1 \leq I \leq m \tag{15}$$

From Equation (14), one can see that the plus sign results in $G_1 > 1$, since

$$2m + I - 1 + \sqrt{4mI + (I-1)^2} > 2m + 2(I-1) > 2(m-1) \tag{16}$$

Furthermore, from Equation (12), one can see that when $0 \leq G_1 \leq 1$, then $0 \leq G_2 \leq m$. Thus, the only realizable maximum is given by Equation (13) with G_1 as in Equation (14) when taking the minus sign of the square root.

Figure 7 shows the system throughput as a function of G_2 for m = 6 and I as a parameter. One can see that there is a high degradation in system performance when the interference level I increases. Figure 8 shows the capacity of the system as a function of the number of repeaters m for the non-interference (I = 1) and the complete interference (I = m) cases. When $1 < I < m$, the curve will be between the two shown. It can be seen that there is a large difference in system capacity between the non-interference and the complete interference systems, and that this difference increases with the number of repeaters m. For I = 1, the capacity of the system is $(m(\sqrt{m} - 1)^2)/(m - 1)^2$ which tends to 1 when m tends to infinity.

In the terrestrial system, the interference level depends on the transmission power of the repeaters when transmitting to terminals. Thus, it would be advantageous to use as low transmission power as possible sufficient to reach the terminals, or possibly an adaptive power mechanism.

A. Directional Antennas and Multiple Transmitters At the Station

This section is addressed only to a terrestrial system. When the station uses a directional antenna then its transmission to repeater R_i does not interfere with R_j, $j \neq i$. Consequently, the average rate of transmission per slot to a single repeater is G_1/m (assuming equal share of load). The only change that would result in Equations (10) and (11) is the replacement of G_1 in Equation (10) by G_1/m. Doing so and equating S_1 and S_2 results in $G_1 = G_2$ and:

$$S_1 = G_1\left(1 - \frac{G_1}{m}\right)^I \; ; \quad 0 \leq G_1 \leq 1 \; ; \; 1 \leq I \leq m \tag{17}$$

$$S_2 = G_2\left(1 - \frac{G_2}{m}\right)^I \; ; \quad 0 \leq G_2 \leq m \; ; \; 1 \leq I \leq m \tag{18}$$

The capacity of the system is given by

$$S_1^* = \begin{cases} \frac{m}{I+1}\left(1 - \frac{1}{I+1}\right)^I ; & m \leq I+1 \quad (19a) \\ \left(1 - \frac{1}{m}\right)^I ; & m > I+1 \quad (19b) \end{cases}$$

Figure 9 shows the capacity of the system as a function of m for I = 1 and I = m. The same curves for an omni-directional antenna are shown as a reference. It can be seen that the capacity of the system with a directional antenna is substantially higher; in particular, when the interference level is low.

We now address the question of multiple transmitters and antennas at the station. Consider the capacity of the system with a directional antenna at the station, Equation (19). The maximum given by (19a) is a stationary point whereas that given by (19b) is a boundary point at $G_1 = 1$. Also note that

$$\frac{m}{I+1}\left(1 - \frac{1}{I+1}\right)^I \geq \left(1 - \frac{1}{m}\right)^I, \text{ for } m \leq I+1 \quad (20)$$

Thus, if one increases the domain of G_1, it would result in an increase in system capacity, which will then be given by Equation (19a). This is exactly what happens when adding additional transmitters and antennas to the station. It enables the station to transmit more than one packet into the same "slot in time" and direct the packets to different repeaters. In practice, if the station has several transmitters and directional antennas which enables it to transmit simultaneously in different directions, then one can

devise an algorithm at the station, which will properly select the directions to which packets are simultaneously transmitted, so as to further reduce the interference level I (manage its transmissions).

Figure 10 shows the system throughput as a function of the rate of transmission from station to repeaters (or from repeaters to terminals, note $G_1 = G_2$). One can see that for I = 1 and I = 3, the maximum system capacity cannot be obtained with a single transmitter and the value obtained is at the boundary point at $G_1 = 1$ and given by (19b). Notice that when the number of repeaters is large and the interference level is low, then there is a large difference between the maximum capacity and the constrained maximum capacity.

We now determine the minimum number of transmitters and directional antennas needed at the station. If the interference level I is constant then the unconstrained capacity of Equation (19a) is increasing with the number of repeaters m. Moreover, the capacity increases also in the case that I is a linear function of m. For let I = km, $1/m \leq k \leq 1$, then the unconstrained capacity is given by

$$S^* = \frac{m}{km + 1} \left(1 - \frac{1}{km + 1}\right)^{km} \tag{21}$$

and

$$\frac{ds^*}{dm} = \frac{km}{(km + 1)^2} \left(1 - \frac{1}{km + 1}\right)^{km - 1} \geq 0; \text{ for } k \geq 0 \tag{22}$$

Since the capacity is increasing as a function of m, one can obtain a capacity greater than 1 (see for example Figure 10, I = 1). The minimum number of transmitters that can

realize the unconstrained capacity is given by the rate of transmission from the station to repeaters at which the maximum utilization is obtained. That is:

$$\text{Minimum Number of Transmitters} = \frac{\lceil m \rceil}{km + 1} = \frac{\lceil m \rceil}{I + I} \qquad (23)$$

where $\lceil x \rceil$ is the smallest integer greater than x. Equation (23) also implies that multiple transmitters will not result in an increase in system capacity when the interference level is high, i.e., $m \leq km + 1$. It is easy to verify that when the constraint on G_1 is satisfied, the constraint on G_2 is also satisfied.

By associating a cost value with a repeater and with each additional transmitter at the station, one can formulate a design optimization problem in which the system cost is traded against the increase in capacity which it results.

The results of this section demonstrate that a single slotted ALOHA channel can be used (and reused) spatially to obtain channel utilization higher than 100%.

5. CONCLUSIONS

Our conslusions from the analysis are outlined.

A. Transmission to Station

1. When the number of repeaters is small (the exact number depends on the interference level), then the critical hop in the network (the capacity bottleneck) is from terminals to repeaters.

2. When the number of repeaters is large, the critical hop is from repeaters to station.

3. The 2-Hop network design which maximizes the system capacity has 2 repeaters when the interference is minimum, and 3 repeaters when the interference is maximum. Additional repeaters reduce system capacity.

4. The capacity of a 2-Hop network is higher than that of a 1-Hop network when the number of repeaters is 2 or more, and is lower than the capacity of a 1-Hop network when there is one repeater.

5. Directional antennas at repeaters increase system capacity when the critical hop is from terminals to repeaters. The increase is significant only in the case when there are 2 repeaters, which do otherwise interfere with each other. In other cases, the increase in capacity is either insignificant or does not exist.

B. Transmission from Station

1. The interference of the station with the transmission of repeaters to terminals reduces significantly the system capacity. Thus, if possible, it is important to enable terminals to receive such transmissions.

2. The system capacity reduces substantially when the interference level between repeaters is increased. Note that this is not the case when transmitting to the station; compare Figures 5 and 8. Consequently, in a terrestrial system, it is important to reduce the interference factor by a mechanism such as adaptive power.

3. A directional antenna at the station in a terrestrial system increases significantly, the system capacity when the interference level between repeaters is low to moderate. This is not the case when the interference level is high, since the throughput on the hop from repeaters to terminals is limited due to this interference.

4. When the station has directional antennas, then multiple transmitters and antennas may further increase, significantly, system capacity. Note that in this case, one can obtain a capacity greater than 1.

5. An equation for the number of transmitters needed at the station is given. This number increases when the interference level decreases. When the interference level is high, then none of the devices in conslusions B.3, B.4, B.5, are desirable, since the capacity is limited by the throughput from repeaters to terminals.

REMARKS:

Conclusions A.5 and B.3 which relate to directional antennas, imply that directional antennas are generally not useful when devices which use them direct transmissions to a single location in space; because the interference at this location is not avoided. On the other hand, directional antennas are very useful when oriented to different azimuths, because one can take advantage of the spatial distribution of receivers.

ACKNOWLEDGEMENT:

The author is grateful to Richard M. Van Slyke for suggesting to consider the case with directional antennas at the station.

REFERENCES

1. Abramson, N., "The ALOHA System - Another Alternative for Computer Communications," Fall Joint Computer Conf. AFIPS Conf. Proc., Vol. 37, 1970, pp. 281-285.

2. Abramson, N., "Packet-Switching With Satellites," National Comp. Conf., June 1973, pp. 695-702.

3. Gitman, I., R. M. Van Slyke, H. Frank, "On Splitting Random Access Broascast Communication Channels," Proceedings of the Seventh Hawaii International Conf. on System Sciences, Subconference on Computer Nets, January 1974.

4. Kleinrock, L. and S. S. Lam, "Packet-Switching in a Slotted Satellite Channel," National Computer Conf., June, 1973, pp. 703-710.

5. Metcalfe, R. M., "Steady-State Analysis of a Slotted and Controlled ALOHA System with Blocking," ARPANET Satellite System Note #16, (NIC Document #11624).

6. Roberts, L. G., ARPANET Satellite System Notes 8, (NIC Document #11291), available from ARPA Network Info. Center, Stanford, Research Institute, Menlo Park, Calif.

7. Roberts, L. G., "Dynamic Allocation of Satellite Capacity Through Packet Reservation," National Computer Conf., June 1973.

8. Kleinrock, L. and S. S. Lam, "On Stability of Packet Switching in a Random Multi-Access Broadcast Channel," Proceedings of the Seventh Hawaii International Conf. on System Sciences, Subconference on Computer Nets, January 1974.

9. Network Analysis Corporation, "Comparison of Hop-by-Hop and End-to-End Acknowledgement Schemes," PRTN #7, Available from ARPA Network Information Center, Stanford Research Institute, Menlo Park, Calif.

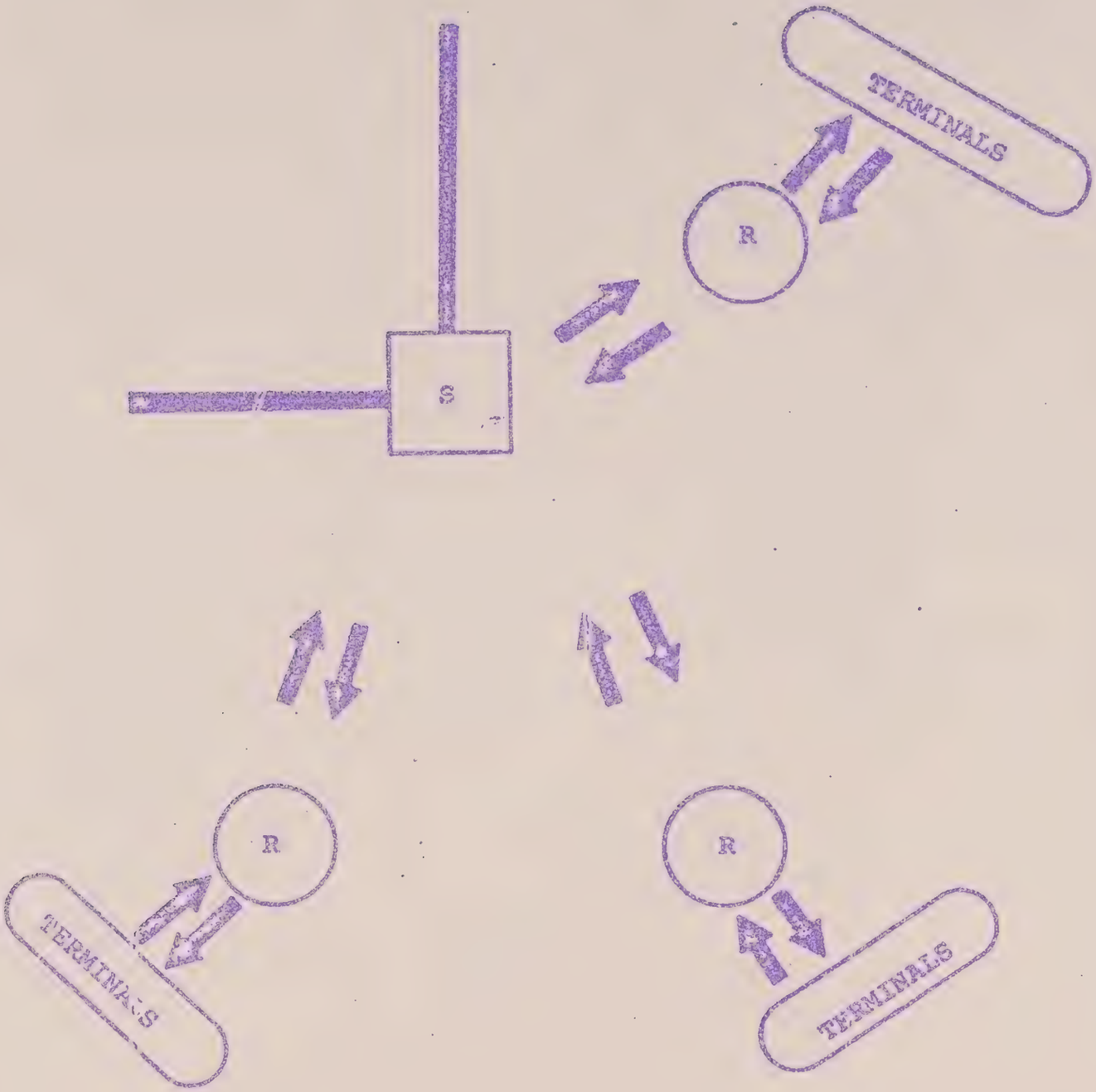


FIGURE 1: A TERRESTRIAL SYSTEM

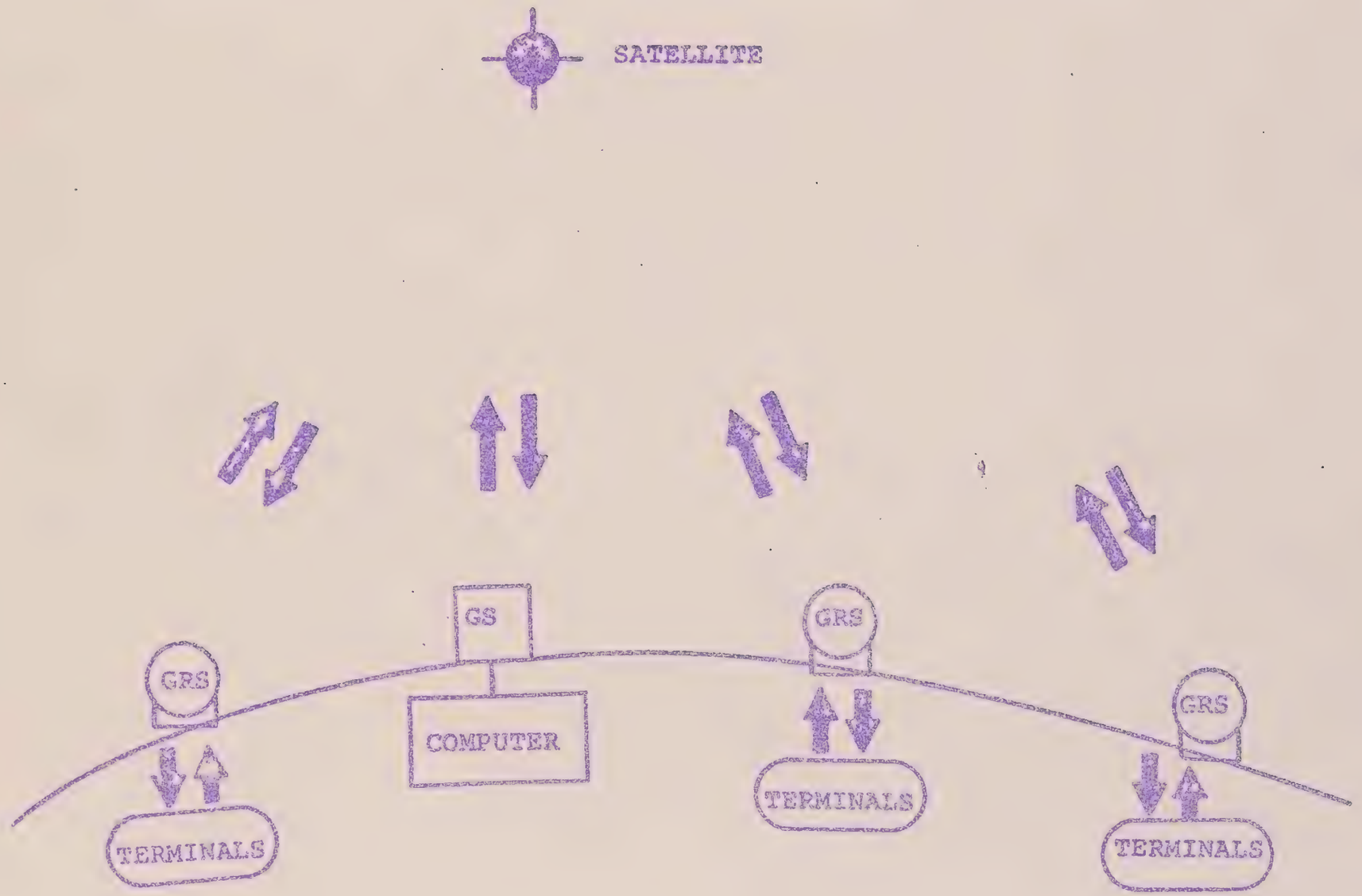


FIGURE 2: A SATELLITE SYSTEM

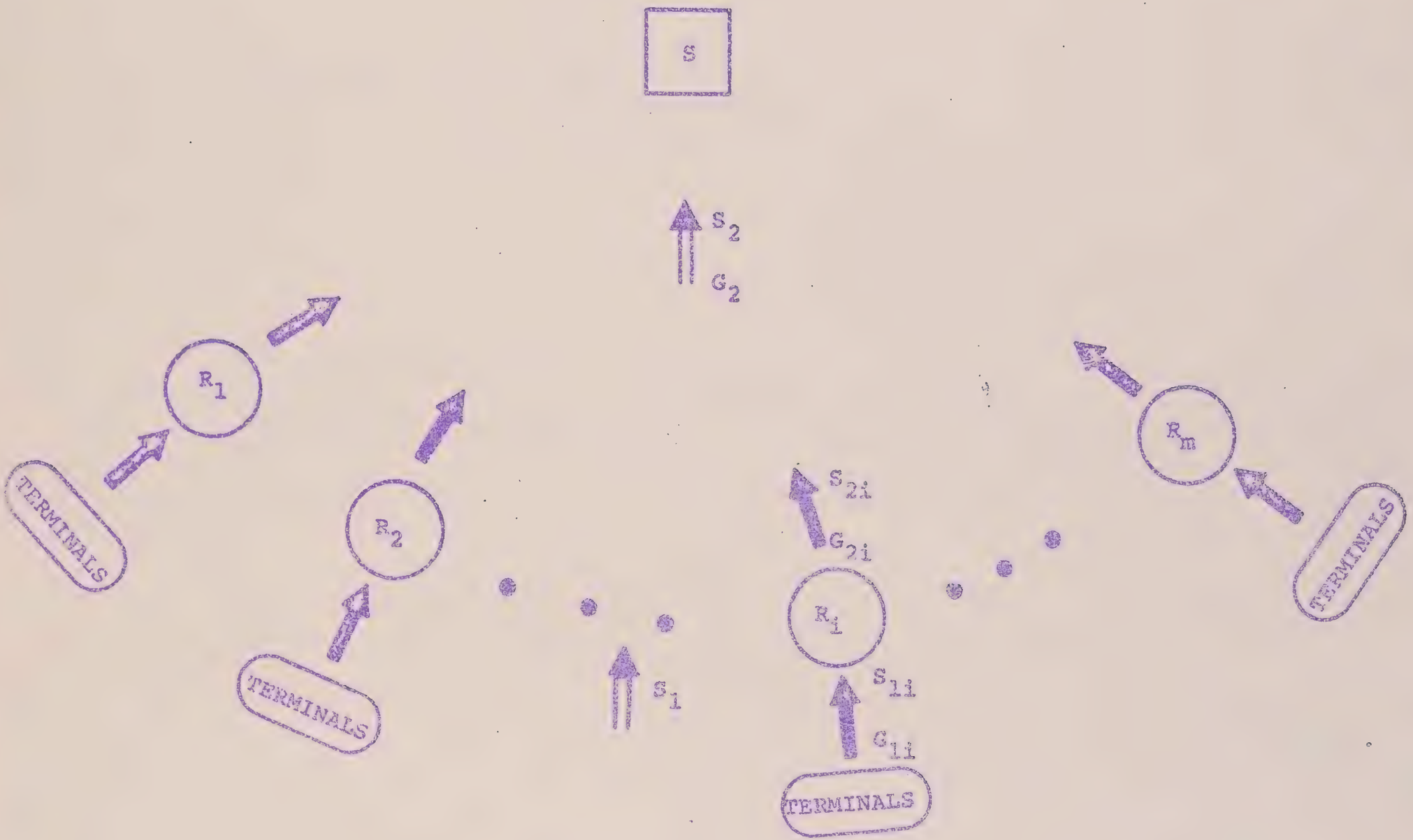


FIGURE 3: TRANSMISSION FROM TERMINALS TO STATION

- 2-Hop Slotted ALOHA Network
- Transmission From Terminals To Station
- Complete Interference System I = m

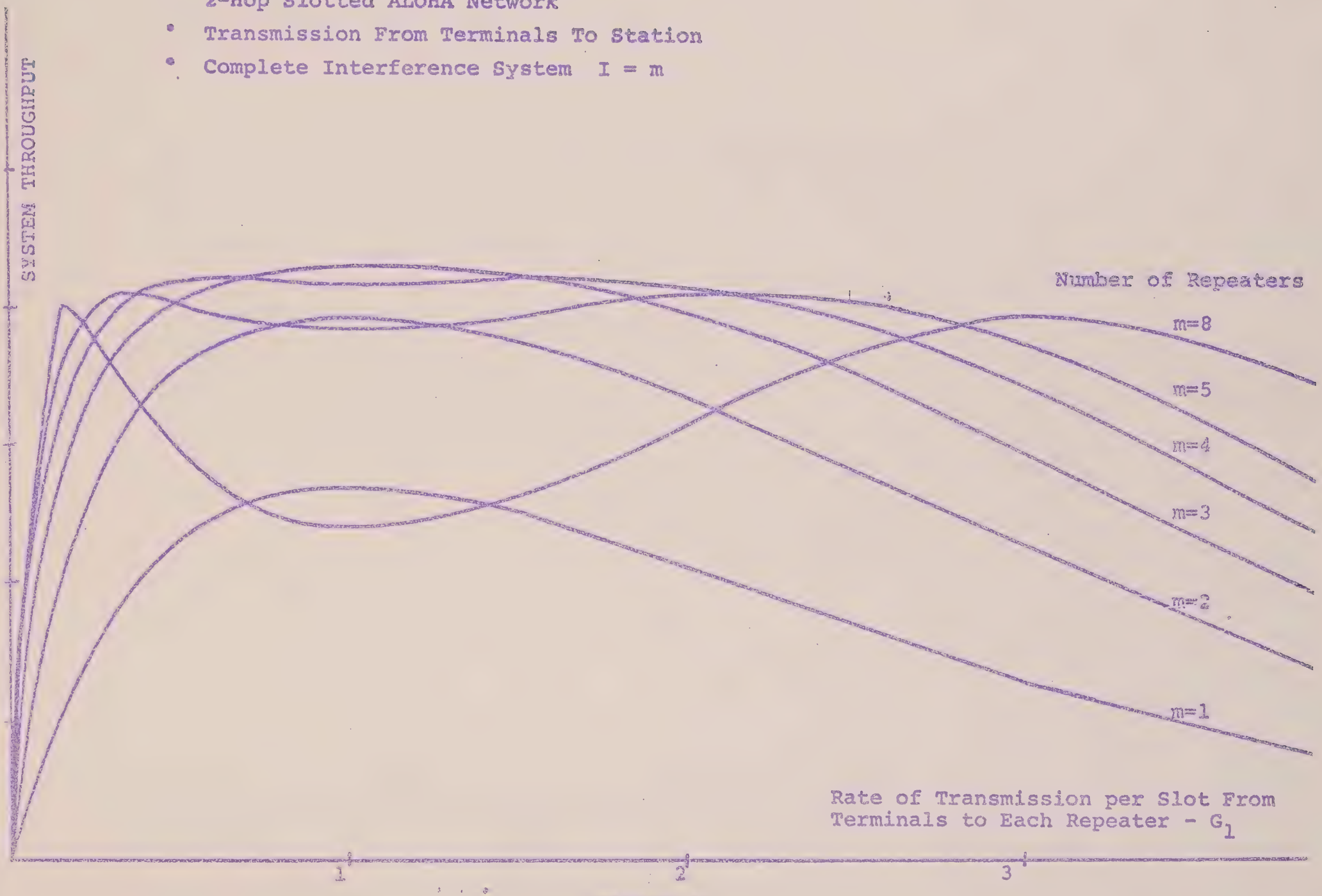


FIGURE 4

- 2-Hop Slotted ALOHA Network
- Transmission From Terminals to Station

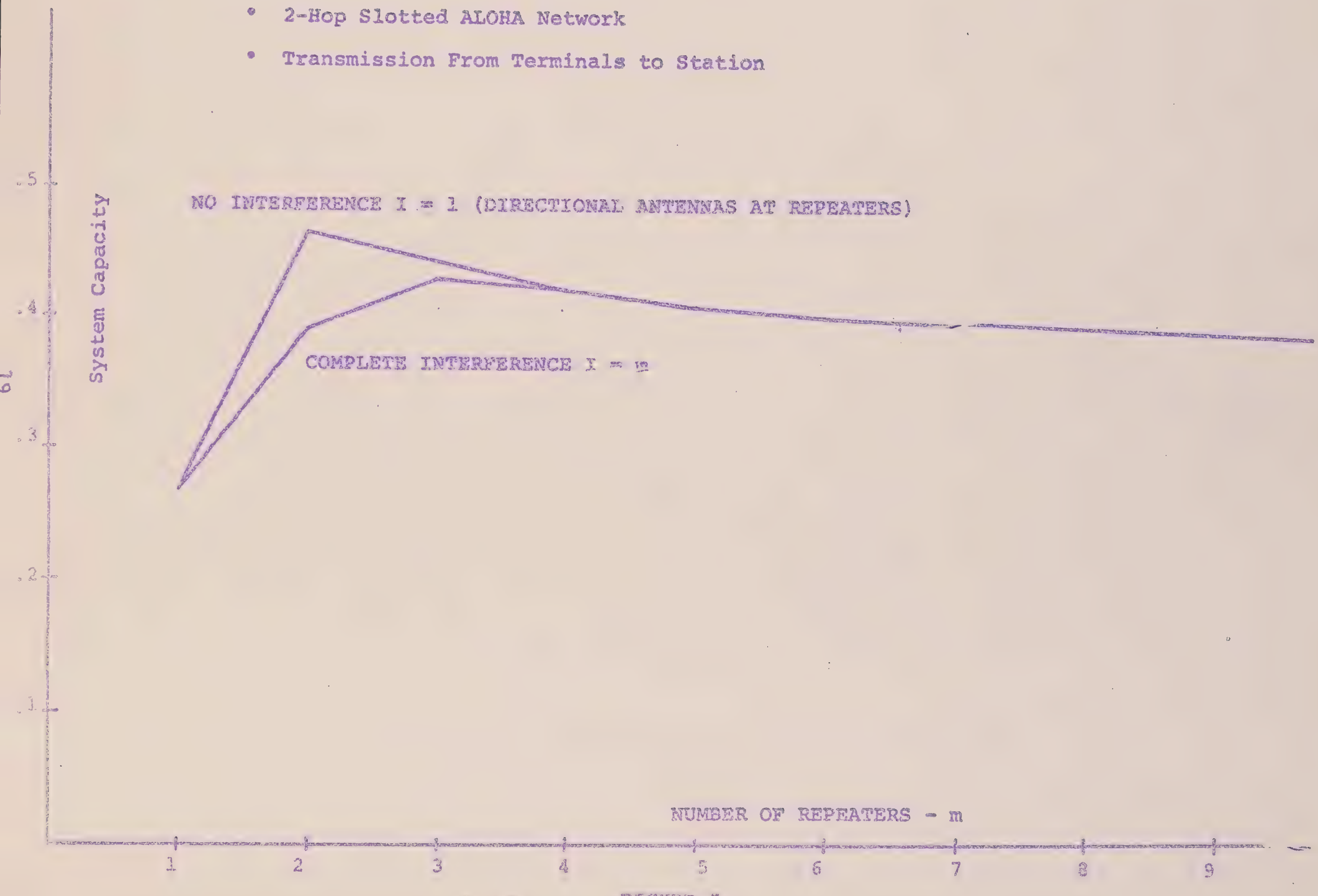


FIGURE 5

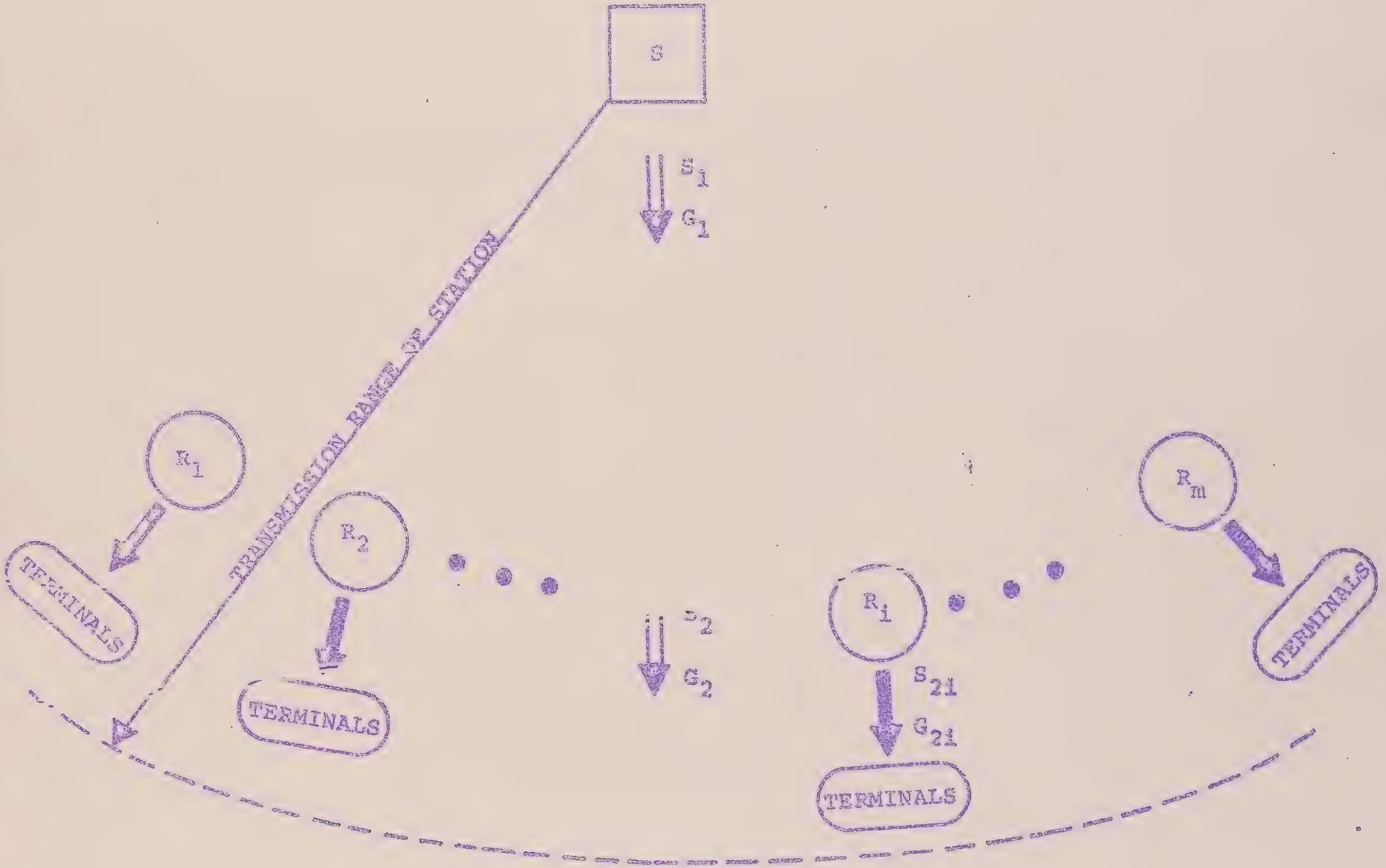


FIGURE 6: TRANSMISSION FROM STATION TO TERMINALS

- 2 - Hop Slotted ALOHA Network
- Transmission From Station to Terminals
- 6 Repeaters

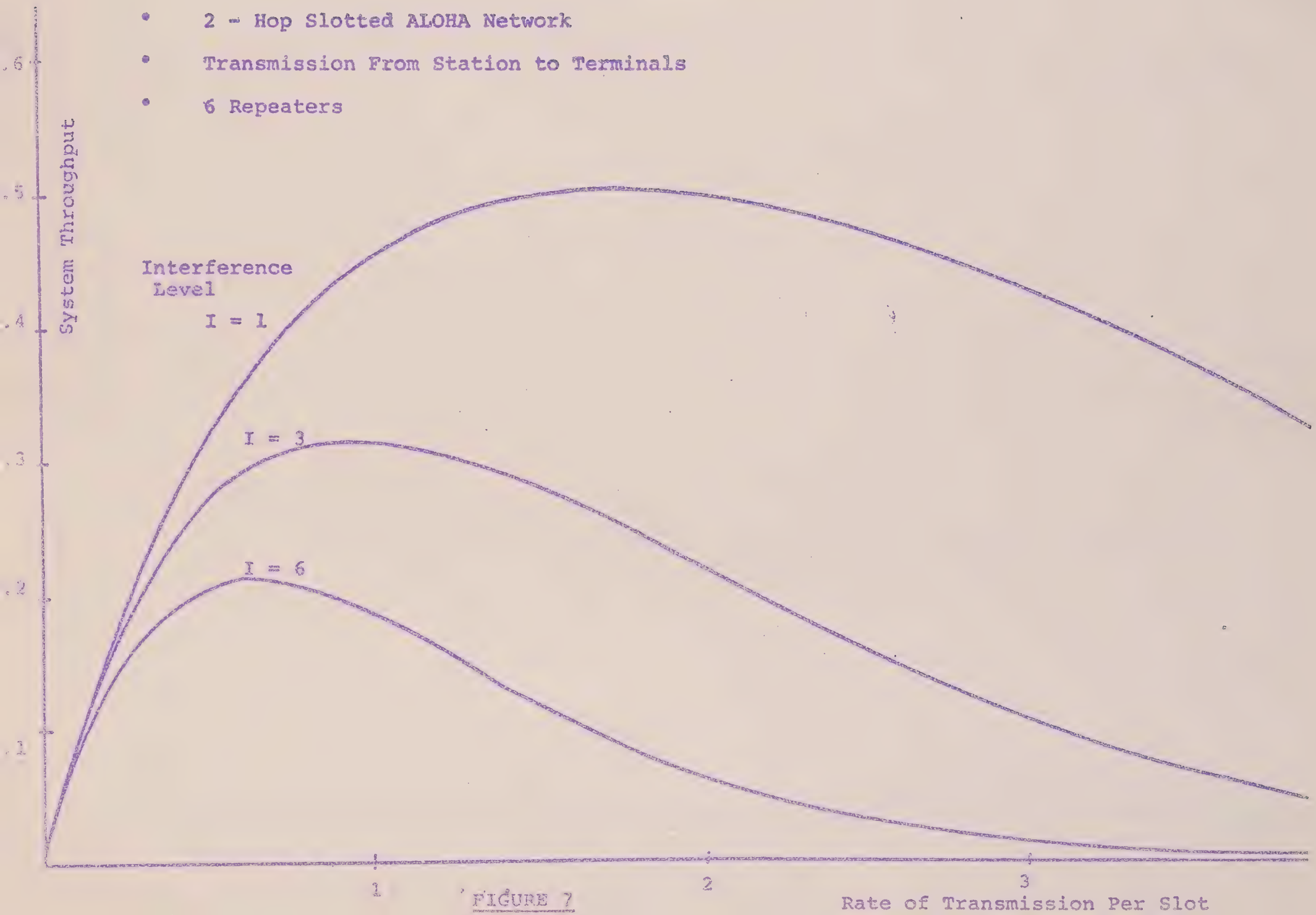


FIGURE 7

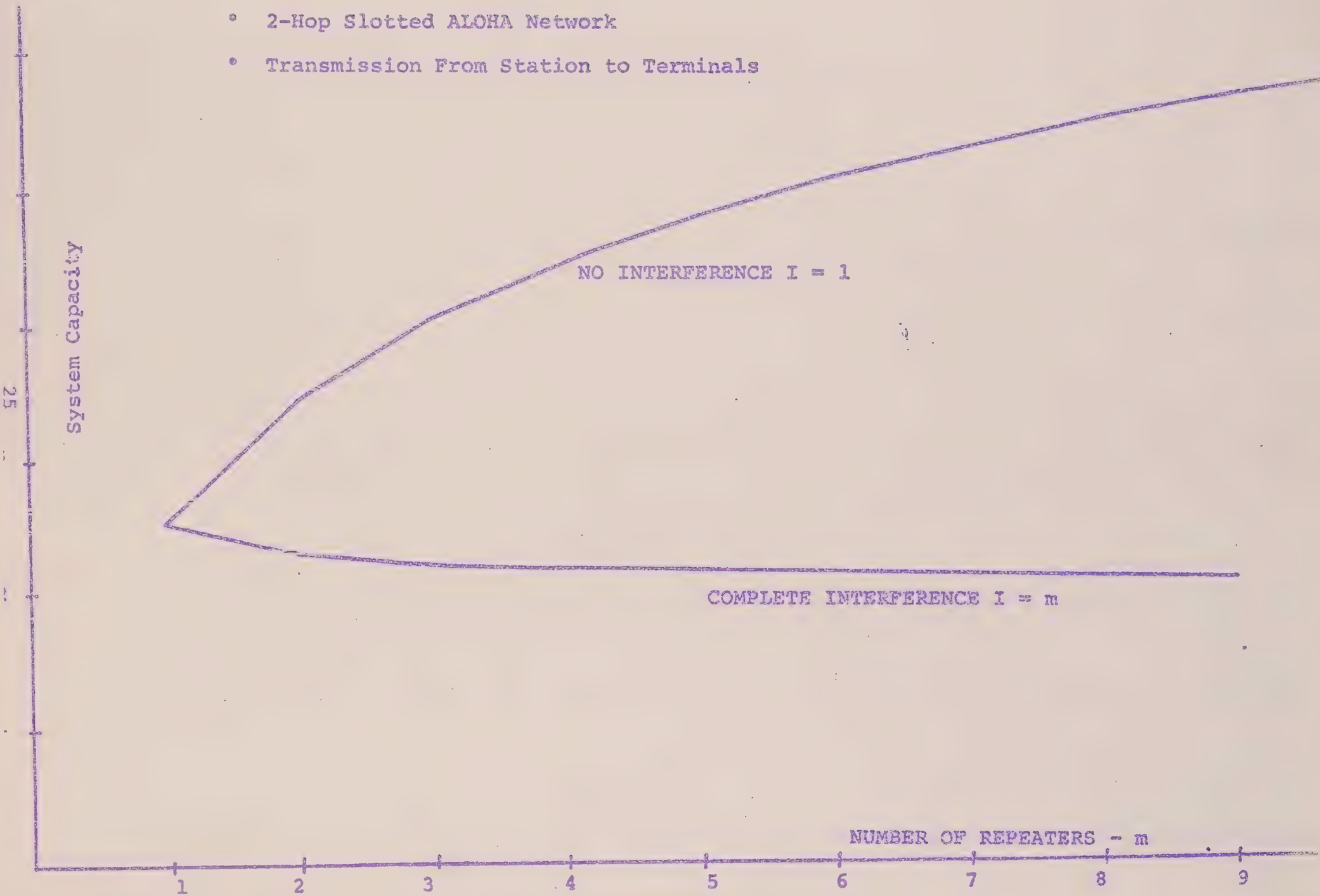

• 2-Hop Slotted ALOHA Network
• Transmission From Station to Terminals
System Capacity
25
NO INTERFERENCE I = 1
COMPLETE INTERFERENCE I = m
NUMBER OF REPEATERS - m
1
2
3
4
5
6
7
8
9

FIGURE 8

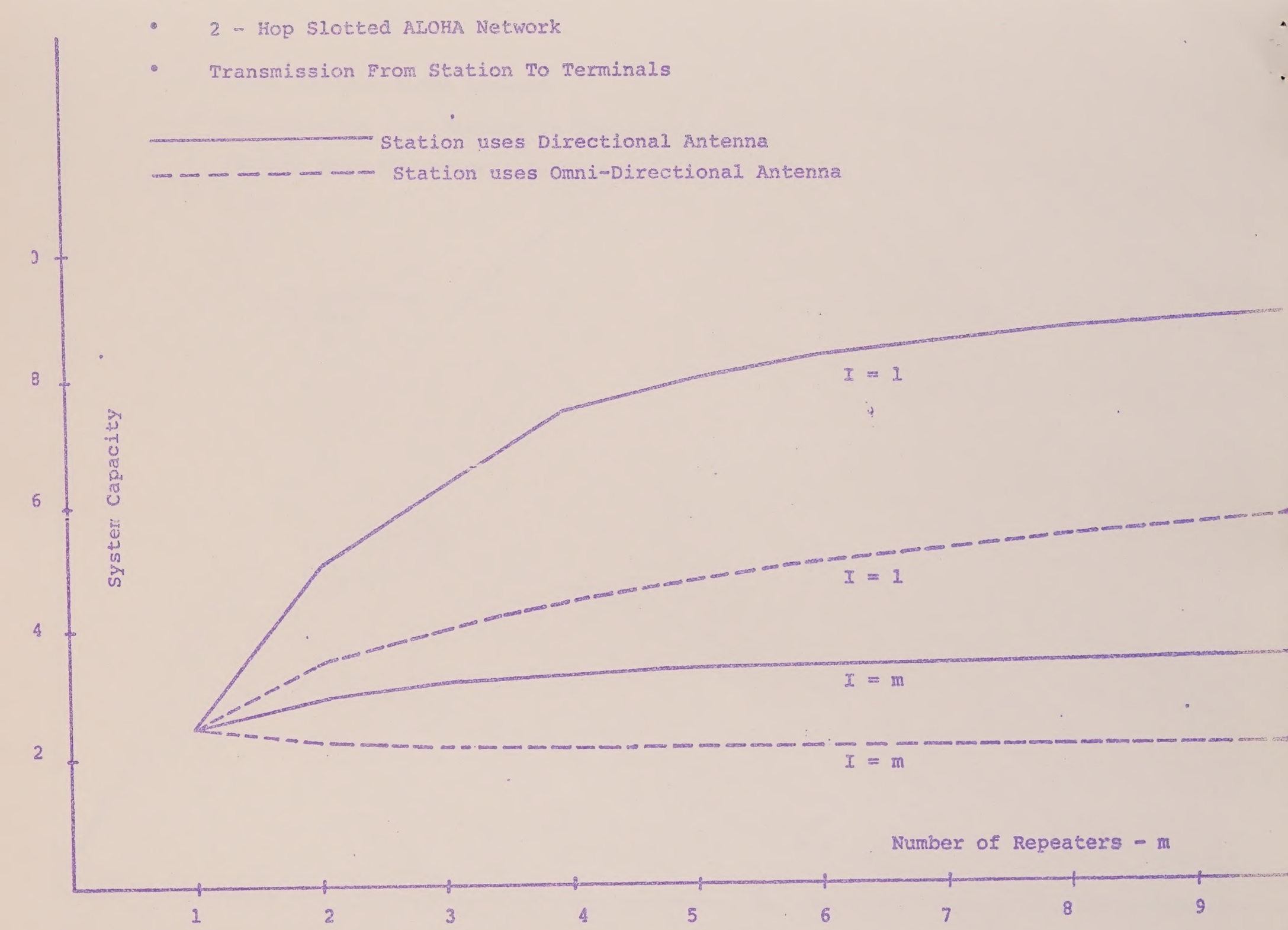
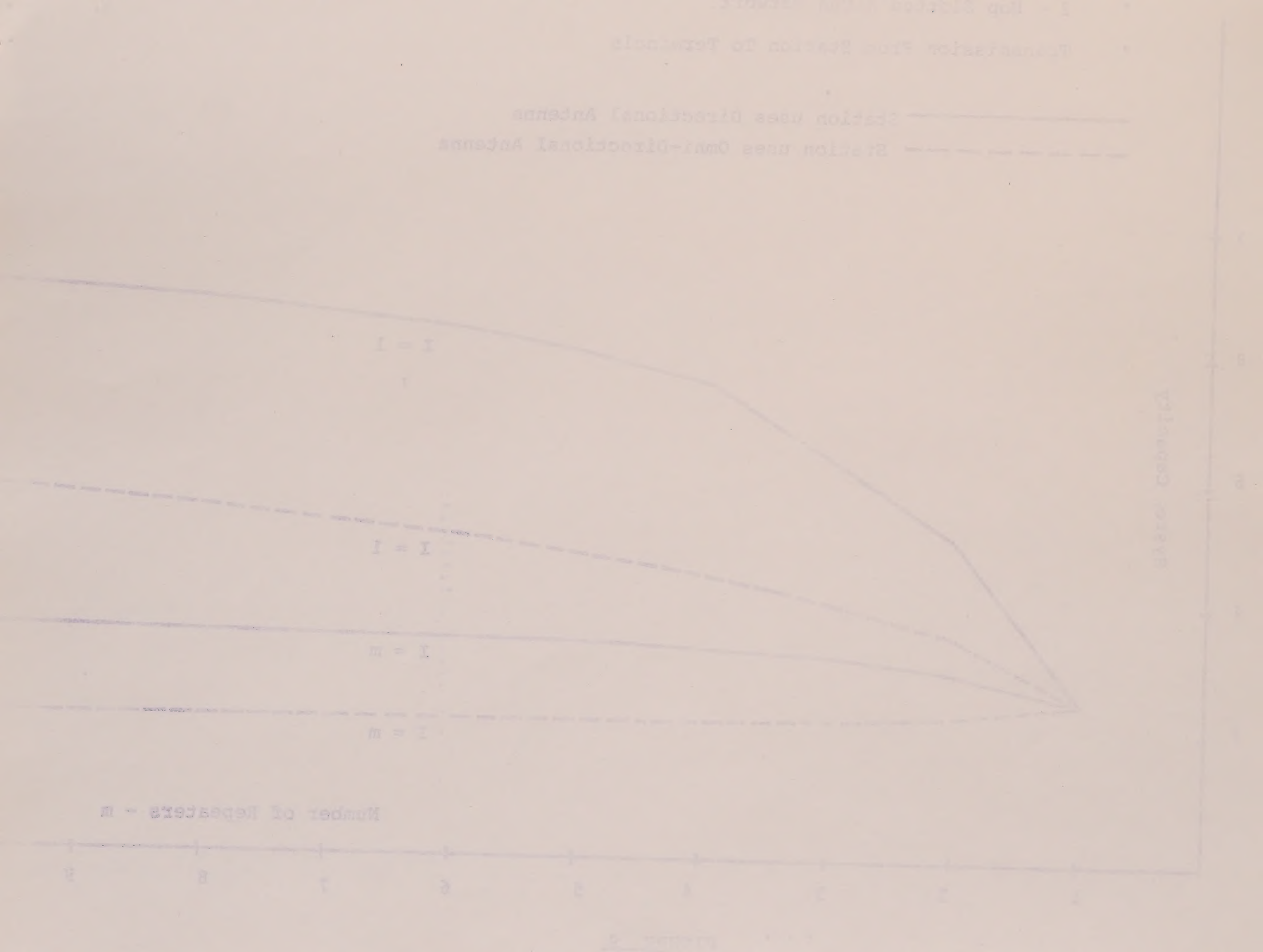
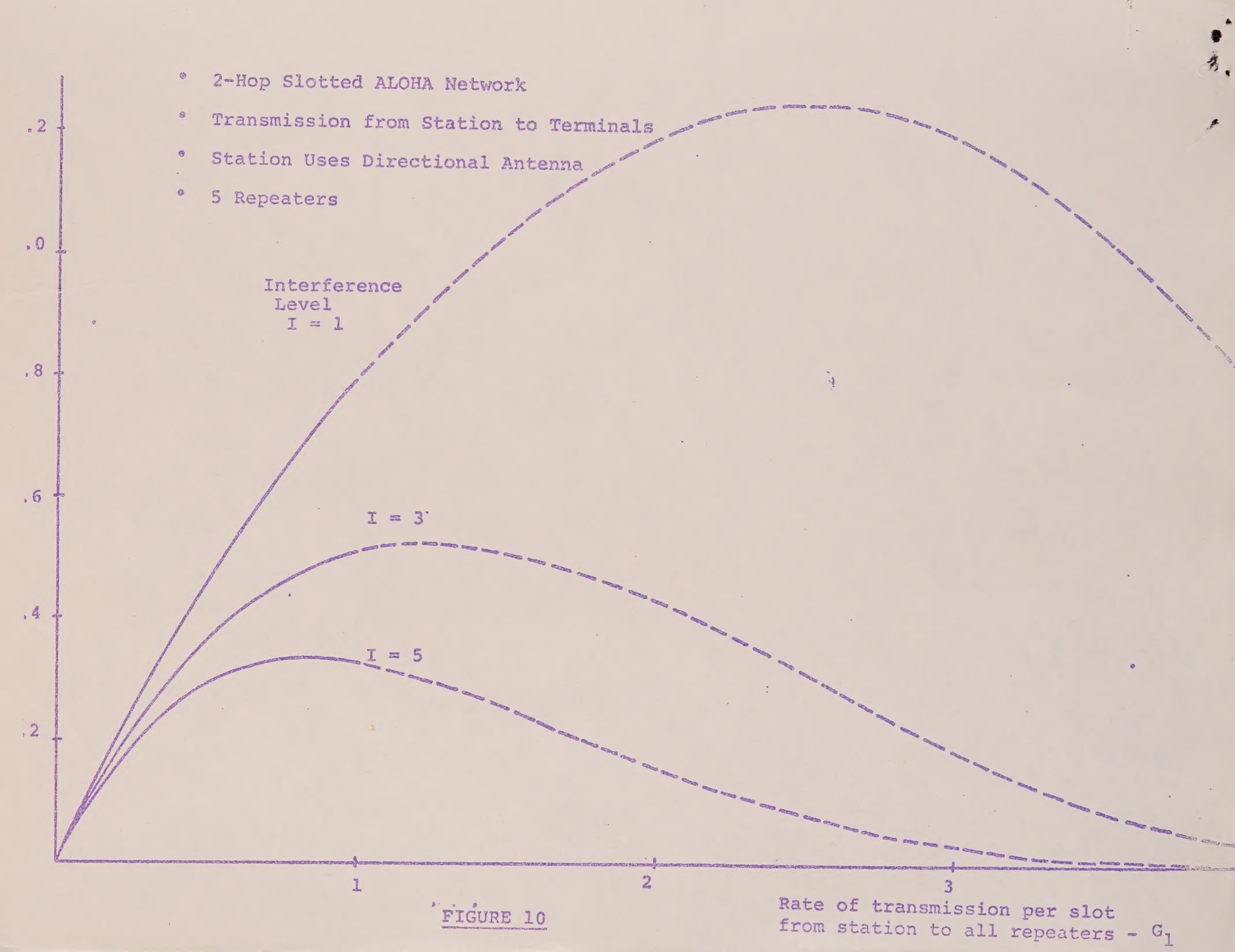

• 2 - Hop Slotted ALOHA Network
• Transmission From Station To Terminals
Station uses Directional Antenna
Station uses Omni-Directional Antenna
System Capacity
I = 1
I = 1
I = m
I = m
Number of Repeaters - m
1
2
3
4
5
6
7
8
9
2
4
6
8

FIGURE 9




2-Hop Slotted ALOHA Network
Transmission from Station to Terminals
Station Uses Directional Antenna
5 Repeaters
Interference Level I = 1
I = 3
I = 5
.2
.0
.8
.6
.4
.2
1
2
3
Rate of transmission per slot from station to all repeaters - G_1

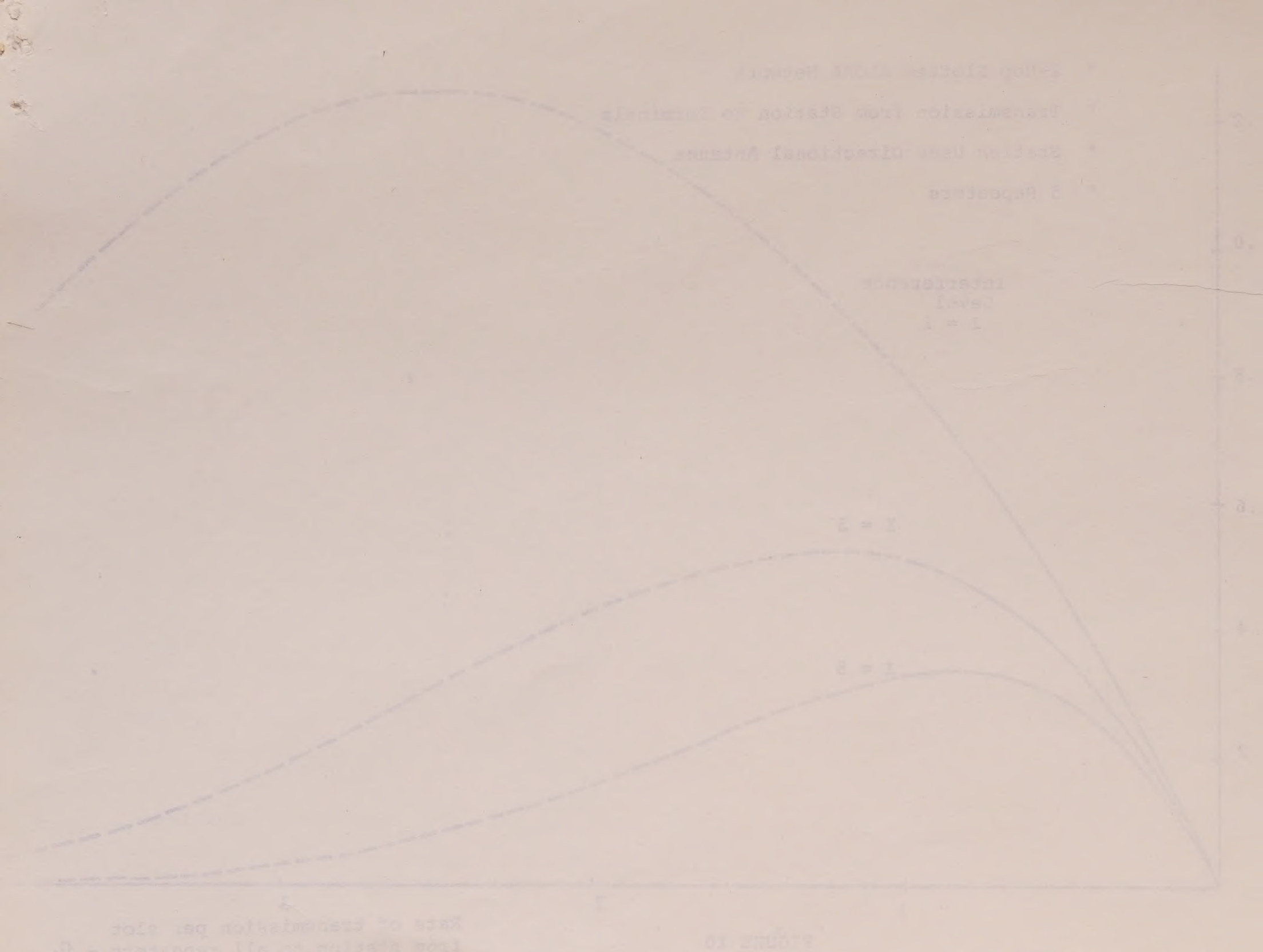



FIGURE 10